DOCUMENT RESUME

ED 397 640 FL 023 851

AUTHOR	Kubota, Mikio
TITLE	Teachability of Conversational Implicature to Japanese EFL Learners.
PUB DATE	95
NOTE	35p.
PUB TYPE	Journal Articles (080) -- Reports - Research/Technical (143)
JOURNAL CIT	IRLT (Institute for Research in Language Teaching) Bulletin; v9 p35-67 1995
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	College Students; *Consciousness Raising; Discourse Analysis; *English (Second Language); Foreign Countries; Higher Education; Japanese; Language Patterns; Learning Strategies; *Pragmatics; Second Language Instruction; *Speech Acts
IDENTIFIERS	*Conversational Analysis; Japan; *Japanese People

ABSTRACT

As teaching pragmatic competence is considered to be one of the neglected aspects in English language teaching in Japan, this paper investigates the teaching of conversational English implicature of 126 Japanese English-as-a-Foreign-Language (EFL) learners. University student participants were divided into three groups and given a multiple choice test and a sentence-combining test. In one group, the explanations of rules were given by a teacher; in the second, consciousness-raising tasks evolved from group discussion. The third group was a control. All subjects received a pre-test and two post-tests. Results indicate that experimental groups generated significantly better responses. In addition, no subjects extracted the expected pragmatic generalizations from the treatment that they were applying to the new items. Also, the conscious-raising groups performed better in the post-test than in the pre-test, and they had significantly higher scores in the guessing of items on the first post-test than the pre-test. Results confirm that teaching conversational implicature through explicit explanations of rules and consciousness-raising tasks is highly facilitative, amount of time and exposure to the pragmatic system may be a crucial factor to induction, and it may be advantageous for learners to process language on their own through consciousness-raising tasks. (Contains 20 references.) (Author/NAV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 397 640

Teachability of Conversational Implicature to Japanese EFL Learners

紀要第9号抜刷

1995, Vol 9

The IRLT Bulletin

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Mikio Kubota

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL023857

財団法人　語学教育研究所

Tokyo The Institute for Research in Language Teaching

BEST COPY AVAILABLE

Teachability of Conversational Implicature to Japanese EFL Learners*

KUBOTA Mikio

ABSTRACT

Teaching pragmatic competence is considered to be one of the neglected aspects in English Language Teaching in Japan. This research is intended to investigate whether it may be effective to teach 'conversational implicature' to Japanese EFL learners.

A total of 126 university students participated in this research. Two types of tests were given to them : a multiple-choice test and a sentence-combining test. The subjects were divided into three groups:

Experimental Groups — [Group A] Explanations of rules given by the teacher
[Group B] Consciousness-raising tasks in group discussion
Control Group — [Group Z] No treatment

The results of the Pre-test, Post-test 1 (immediate post-test), and Post-test 2 (post-test given one month after treatment) indicate that:

(1) all the experimental groups generated significantly better responses than the control group, at least temporarily,
(2) the subjects performing consciousness-raising tasks (Group B) performed significantly better in Post-test 1 than in the Pre-test,
(3) no subjects extracted the expected pragmatic generalizations from the treatment that they were applying to the new items,
(4) the subjects in Group B gained significantly higher scores in the guessing items of Post-test 1 than the Pre-test.

The pedagogical implications that teachers should keep in mind are:

(1) teaching conversational implicature through explicit explanations of rules and consciousness-raising tasks to EFL learners in Japan is highly facilitative,

(2) it may be advantageous for learners to process language on their own through consciousness-raising tasks,

(3) the amount of time and exposure to the pragmatic system may be a crucial factor to induction,

(4) the effect of Pre-test experiences may influence the results.

1. INTRODUCTION

The notion of 'conversational implicature' (Grice 1975) is significant in recognizing and conveying a message in a conversation. The listener tries to search for another possible meaning of the message that the speaker intends to convey, when the literal meaning is not what the speaker intends.

Grice (1975) proposed a cooperative principle which has a set of the following four conversational maxims:

1. The maxim of quantity
   —Make your contribution as informative as is required.
   —Do not make your contribution more informative than is required.
2. The maxim of quality
   —Do not say what you believe to be false.
   —Do not say that for which you lack adequate evidence.
3. The maxim of relevance
   —Be relevant.
4. The maxim of manner
   —Be brief and orderly.
   —Avoid obscurity of expression and ambiguity.

Conversational implicature may follow and/or break these conversational maxims. For instance,

A: Where's the fish?

B: *The cat looks happy*. (cf. Kinoshita 1992:149)

This dialogue, where B implicates that the cat probably ate it, is examined in terms of four conversational maxims:

1. The maxim of quantity

This sentence is not informative at all for the current purpose of the exchange, since B does not directly tell the location of the fish.

2. The maxim of quality

This maxim is acceptable in this dialogue, since it seems that B utters the true information of the scene that B watched.

3. The maxim of relevance

It appears that the location of the 'fish' in A's utterance and the feeling of the 'cat' in B's utterance are not related. B's utterance never becomes the direct answer to A's question. Therefore, the dialogue breaks the maxim of relevance.

4. The maxim of manner

The dialogue follows the maxim of manner, because it avoids unnecessary prolixity and ambiguity. Hence, the dialogue breaks the maxims of quantity and relevance.

There were few studies regarding the teachability of pragmatic knowledge[1]. Cohen and Olshtain (1988, as cited in Billmyer 1990:31) showed that there existed positive effects for instruction in the rules of apologizing on written tests in class.

Wilder-Bassett (1984) studied acquisition of gambits (routines for conversational management and modification of illocutionary force) in a classroom setting. She found that learners' use of gambits improved significantly, qualitatively and quantitatively, regardless of teaching approach. However, learners taught according to a vaguely communicative approach were even more successful than their colleagues who had been exposed to a version of suggestopedia (Wilder-Bassett 1984, as cited in Kasper 1994:31).

Billmyer (1990) investigated the production of compliments and replies

to compliments by two different groups of 18 Japanese female ESL learners in America (intermediate and advanced levels of proficiency) during social interactions with native speakers of English outside of the classroom. One group received formal instruction—six hours of supplemental instruction biased toward the explicit presentation of the rules of complimenting in American English. The other group was not given this additional instruction, although the untutored subjects, as well as the tutored subjects, were enrolled in regular ESL courses. All subjects in both groups were asked to perform certain compliment-inducing tasks, such as showing photos of their home and family members, reporting on an accomplishment, and showing a recently purchased item of apparel. Billmyer (1990:44) concluded that on the measures of frequency of compliments, level of spontaneity, adjectival repertoire, reply types and its length, the tutored group exhibited behavior more closely approximating native speaker norms in complimenting than the untutored group. These findings lend support to the hypothesis that formal instruction concerning the social rules of language use given in the classroom can assist learners in communicating more appropriately with native speakers of the target language in meaningful social interaction outside of the classroom (Billmyer 1990:44).

Bouton (1988) made a cross-cultural study of the ability of nonnative English-speaking (NNS) university students to interpret implicatures in English, finding that NNS arrived at the same interpretation of the implicatures as the American NS only about 79% of the time.

Bouton (1994) longitudinally examined two groups of international students at an American university with regard to their ability to interpret implicature, when it was not taught deliberately. The first group of 30 subjects took part in the multiple-choice test (referred to as the implicature test) for the 4½-year study (i.e., the original study—in 1986 and the second one—in 1991). No significant difference was found in the scores on 20 items of the implicature test in 1991 between the American

native speakers (NS) and NNS subjects ($p > .3056$). By contrast, a comparison of the scores of the NS and NNS in 1986 produced a significant difference ($p < .0001$). Hence, it was clear that the scope of the difference between the NS and the NNS diminished over the NNS subjects' 4½ years of residence. It was also found that the number of items that were interpreted differently was greatly reduced and none of the types of implicature that were troublesome for the NNS in 1986 were consistently causing trouble in 1991.

The second group of 34 NNS subjects were given a modified version of the implicature test in the 17-month study (in 1990 and in 1992). The result shows that the subjects responded significantly better in 1992 than in 1990 ($p < .0001$). However, the scores of the NNS and the NS were significantly different ($p < .019$). Accordingly, it was concluded that the NNS subjects improved their proficiencies in the interpretation of implicature, although they were not yet native-like. Most significant, they mastered none of the types of implicature that bothered them initially 17 months before. Comparing the results of these two groups, Bouton (1994) concluded that the process of learning conversational implicature is slow when implicature is not deliberately taught.

ESL textbook materials include relatively few examples of conversational implicature and few of those are dealt with directly (Bouton 1990). Furthermore, EFL textbooks approved by the Ministry of Education in Japan do not deal with conversational implicature as a whole, either.

In particular, no research has been conducted in the teachability of conversational implicature in ESL/EFL situations to the best of the present researcher's knowledge. This research will, thus, be a pioneer experiment conducted in Japan.

2. THE STUDY

2.1. Research Questions

The main purpose of this research is to examine whether conversa-

tional implicature will be teachable to Japanese university students of EFL. The following two Research Questions are posed:

Research Question (1) — What types of treatment will facilitate learners' pragmatic competence?

Research Question (2) — How effective will treatment be on a long-term basis?

2.2. Hypotheses

Hypotheses 1-3 are concerned with Research Question (1), whereas Hypotheses 4-7 are related to Research Question (2).

H1: There would be no statistically significant difference in the accuracy of responses between the experimental groups and the control group.

The test scores would result in no statistically significant difference between the experimental groups and the control group. If this null hypothesis proves incorrect, the alternative hypothesis is stated as follows:

H2: The experimental groups would outperform the control group in Post-test 1 (immediate post test) and Post test 2 (delayed post-test).

It is hypothesized that providing feedback would be influential in inducing a positive learning effect. Previous researches (Carroll et al. 1992, Carroll and Swain 1993, Kubota 1994)[2] reveal that the experimental groups which received corrective feedback outperformed the comparison groups which received no feedback. Thus, these studies led to the formation of Hypothesis 2.

H3: There would be no statistically significant difference in the accuracy of responses between the experimental groups.

There are two types of experimental groups in this study: one type is the group that receives explanations of rules regarding conversational implicature explicitly, and the other type is the group that performs consciousness-raising tasks. This is a null hypothesis, since there is no

previous research or theory that explains the difference.

H4: There would be no statistically significant difference in the accuracy of responses among the Pre-test, Post-test 1, and Post-test 2, in the experimental groups.

This is also a null hypothesis, since no theory or previous research[3] clearly explains the difference. The following alternative hypotheses 5-7 about the relationship between instruction and L2 learning will be tested in case the null hypothesis is rejected.

H5: The experimental groups would respond significantly better in Post-test 1 than in the Pre-test.

H6: The experimental groups would respond significantly better in Post-test 2 than in the Pre-test.

It is assumed that a positive effect of instruction on learning would appear in the experimental groups. Hence, the subjects receiving explanations of the rules and the subjects performing consciousness-raising tasks would respectively get higher scores in the Post-tests 1 and 2 than in the Pre-test.

H7: The experimental groups would respond significantly better in Post-test 1 than in Post-test 2.

There was no class conducted between the treatment and Post-test 2, owing to the winter vacation, and all the subjects were asked not to study conversational implicatures outside class. Accordingly, it seems that they had no opportunities to hear and use the target points. It is hypothesized that the effect of instruction would begin to disappear.

2.3. Data collection and analysis

2.3.1. Subjects

A total of 126 Japanese university students participated in this research. They had studied EFL for six or seven years in only instructional settings. They reported that they had had no experiences in studying conversational implicatures, while they were in junior and

senior high schools.

Intact classes were used in this research, because it was impossible to assign students randomly to the experimental and control groups. 35 students had to be excluded from the analysis, since they missed one or more of the following: the Pre-test, the treatment, Post-test 1, or Post-test 2.

2.3.2. Test Items

There were two tests given to all the subjects: Test (A) and Test (B). Test (A) was a multiple-choice test, while Test (B) was a sentence-composing test. Test items in all three sessions (Pre-test session, Post-test 1 session, and Post-test 2 session) were all the same (see Appendix).

[Test (A)]

The subjects were asked to choose the most appropriate answer in the multiple-choice test. The following four types of conversational implicatures were tested in Test (A) which had six test items (cf. Grice 1975, Bouton 1994; see Appendix):

(1) Sequence ·····No. 1 & No. 4 in Test (A)
(2) POPE Q implicature[4]·····No. 2 & No. 6 in Test (A)
(3) The maxim of relevance ···No. 3 in Test (A)
(4) Understated criticism ·····No. 5 in Test (A)

[Test (B)]

The subjects were required to write the conversational implicature of the speaker in English. The following five types of conversational implicatures were tested in Test (B) that had ten test items (cf. Grice 1975, Bouton 1994; see Appendix):

(1) Not breaking any maxim ·····No. 1 & No. 5 in Test (B)
(2) Breaking the maxim of quantity·····No. 2 & No. 7 in Test (B)
(3) Breaking the maxim of quality ·····No. 3 & No. 9 in Test (B)
(4) Breaking the maxim of relevance ···No. 4 & No. 10 in Test (B)

(5) Breaking the maxim of manner ……No. 6 & No. 8 in Test (B)

2.3.3. Procedures

[Step 1: Pre-test]

All the subjects in six classes were given a 30-minute Pre-test (see Appendix).

[Step 2: Treatment]

Six classes were randomly assigned to the following groups: two experimental Groups A, two experimental Groups B, and two control Groups Z:

Experimental groups: Group A—Explanations of rules [42 subjects]
Group B—Consciousness-raising tasks [42 subjects]
Control group: Group Z—No treatment [42 subjects]

The subjects in the experimental groups received treatment in half of all the test items, so that they had to guess the correct responses in the other half: Test (A)—the 'feedback' items were Nos.1, 2, 3, while the 'guessing items' were Nos.4, 5, 6; Test (B)—the 'feedback' items were Nos. 1, 2, 3, 6, 10, while the 'guessing items' were Nos.4, 5, 7, 8, 9.

Group A: The subjects in Group A were given an explicit explanation of the rules concerning conversational implicature, and the answers by the teacher.

Group B: The subjects in Group B did consciousness-raising tasks in small groups. Ellis (1994:643) refers to consciousness-raising tasks as tasks in which "the learners are not expected to produce the target structure, only to understand it by formulating some kind of cognitive representation of how it works, and as tasks directed only at explicit knowledge." In this research, the subjects in Group B were requested to have a four-member[5] group discussion in English to find and reconfirm the rules of conversational implicatures and the answers by themselves, by raising their

consciousness of the target points. After the group discussion, the subjects in Group B were told the correct answers by the teacher. This consciousness-raising task combines the development of explicit knowledge about problematic L2 features with the provision for meaning-focused use of the target language.[6] Accordingly, the difference of treatments between Group A and Group B was that in Group A the subjects received explanations provided by the teacher, whereas in Group B they shared information with each other.

Group Z: The subjects in Group Z, the control group, received no treatment.

Each group consisted of 42 subjects. Only one Japanese teacher of EFL was selected for the purpose of controlling variables (treatment time, treatment contents), and treatment was provided to each group collectively in Japanese, 20 minutes after the Pre-test during the same class period. The reason why there was a 20-minute intermission was that this research was designed to minimize test fatigue of the subjects. Each experimental group took 20 minutes to receive treatment.

[Step 3: Post-test 1 (immediately after treatment)]

Post-test 1 was given 20 minutes after treatment. It took 30 minutes during the same class period as the Pre-test.

[Step 4: Post-test 2 (one month after treatment)]

The 30-minute Post-test 2 was given to the subjects one month after Post-test 1.

2.3.4. Data Analysis

A level of significance ($\alpha = .05$) was selected. The z test and the two-way repeated-measures ANOVA were employed to analyze the data.

3. RESULTS

3.1. Test (A)

3.1.1. Feedback Items—Test (A)

The full mark of the feedback items in Test (A) was 3 points. Table 1 shows the means and standard deviations of the experimental groups and the control group. The means of correct responses for feedback items in Test (A) are displayed in Figure 1.

Table 2 shows the results of the two-way repeated-measures ANOVA for the feedback items in Test (A). The results indicate that group differences and session differences were statistically significant ($p < .05$), as shown in Table 2. Furthermore, the group by session interaction was statistically significant. Hence, the main effect for the group, the main effect for the session, and the effect for the group by session interaction were observed, so that groups and sessions influenced the results dependently.

Table 1 : Means and standard deviations by group and session for the feedback items in Test (A)

Group	n	Mean	Standard Deviation
[Pre-test]			
A	42	1.81	0.82
B	42	1.57	0.73
Z	42	1.50	0.70
[Post-test 1]			
A	42	2.83	0.53
B	42	2.88	0.39
Z	42	1.40	0.79
[Post-test 2]			
A	42	2.43	0.76
B	42	1.86	0.77
Z	42	1.31	0.88

Figure 1 : Means of correct responses for the feedback items in Test (A)

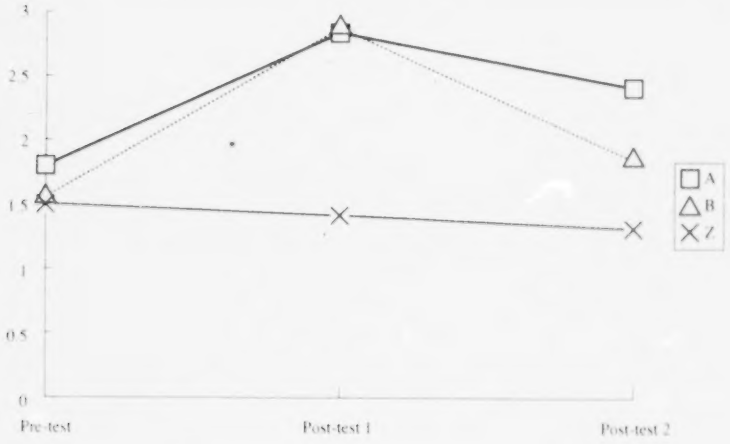

Table 2 : Results of two-way repeated-measures ANOVA for the feedback items in Test (A)

Source	SS	df	MS	F	p
Between subjects	167.57	125			
Groups	61.30	2	30.65	35.64	<.05
Subjects within groups	106.27	123	0.86		
Within subjects	148.67	252			
Session	36.60	2	18.30	52.29	<.05
Groups by session	26.32	4	6.58	18.80	<.05
Residual	85.75	246	0.35		

$F_{2,123}(.05) = 3.09$
$F_{2,246}(.05) = 3.04$
$F_{4,246}(.05) = 2.41$

Table 3 : Results of two-way repeated-measures ANOVA of the interaction effects for the feedback items in Test (A)

Source	SS	df	MS	F	p
Between subjects					
Groups in Pre-test	2.20	2	1.10	0.25	ns
Groups in Post-test 1	59.11	2	29.56	6.66	$<.05$
Groups in Post-test 2	26.31	2	13.16	2.96	ns
Within cell	1636.78	369	4.44		
Within subjects					
Tests in Group A	22.33	2	11.17	31.91	$<.05$
Tests in Group B	39.83	2	19.92	56.91	$<.05$
Tests in Group Z	0.76	2	0.38	1.09	ns
Tests × S	85.75	246	0.35		

$F_{2,369}(.05) = 3.04$

$F_{2,246}(.05) = 3.04$

The interaction effects were tested in order to determine which factors differed from each other. Table 3 shows the results of the two-way repeated-measures ANOVA of the interaction effects for the feedback items in Test (A). It was confirmed that there was no statistically significant difference among groups in the Pre-test; otherwise further analyses would not have been made. Therefore, any comparative effects due to treatment were not related to prior knowledge or language ability of any one group. As shown in Table 3, groups in Post-test 1, the tests in Group A, and the tests in Group B were statistically significant ($p<.05$). Furthermore, multiple comparisons were made in order to decide which factors were different from each other, using Tukey's method. The results of between-group comparisons are shown in Table 4, while Table 5 displays the results of between-session comparisons. As displayed in Table 4, Groups A and B were statistically different from Group Z, respectively, in Post-test 1. Accordingly, it is concluded that treatments given to Groups A and B were more effective than no treat-

BEST COPY AVAILABLE

Table 4 : Between-group comparisons of means for the feedback items of Post-test 1 in Test (A)

Mean	Group	Z	A
1.40	Z		
2.83	A	*	
2.88	B	*	

*p<.05

Table 5 : Between-session comparisons of means for the feedback items in Test (A)

Mean	Session	Pre-test	Post-test 1
[Group A]			
1.81	Pre-test		
2.83	Post-test 1		
2.43	Post-test 2		
[Group B]			
1.57	Pre-test		
2.88	Post-test 1	*	
1.86	Post-test 2		

*p<.05

ment (Group Z), at least for a short-term. The null hypothesis (H1) was rejected, thereby H2 was partially supported.

There was no statistically significant difference found between Groups A and B, in Post-tests 1 and 2, so the data supported the null hypothesis (H3). As illustrated in Table 5, no tests differed from each other in Group A, but Post-test 1 was statistically different from the Pre-test in Group B. Accordingly, in Test (A) H4 was supported, so that H5–H7 did not require testing. On the contrary, in Test (B), H4 was rejected, thereby H5 was supported, although H6–H7 were not supported by Test

(B).

3.1.2. Guessing Items—Test (A)

The full mark in the guessing items of Test (A) was 3 points. The following Table 6 indicates the means and standard deviations by group and session for the guessing items in Test (A). The means of correct responses for the guessing items in Test (A) are illustrated in Figure 2.

Table 7 shows the results of the two-way repeated-measures ANOVA for the guessing items in Test (A). As shown in Table 7, group differences and session differences were not significant. Moreover, the group by session interaction was not significant, either.

Multiple comparisons (Tukey's method) were not made to determine which groups significantly differed from each other, since group differences were not of consequence. Therefore, the null hypotheses (H1 and H3) were supported.

Table 6 : Means and standard deviations by group and session for the guessing items in Test (A)

Group	n	Mean	Standard Deviation
[Pre-test]			
A	42	2.38	0.84
B	42	2.48	0.73
Z	42	2.30	0.64
[Post-test 1]			
A	42	2.62	0.69
B	42	2.48	0.63
Z	42	2.33	0.68
[Post-test 2]			
A	42	2.50	0.73
B	42	2.45	0.79
Z	42	2.48	0.73

Figure 2 : Means of correct responses for the guessing items in Test (A)

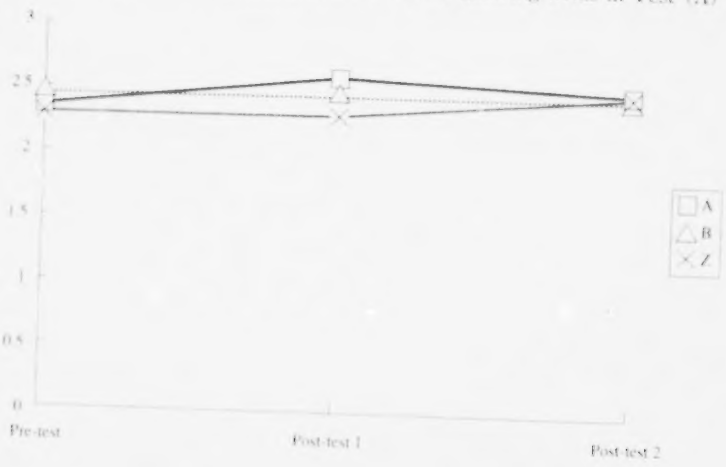

Table 7 : Results of two-way repeated-measures ANOVA for the guessing items in Test (A)

Source	SS	df	MS	F	p
Between subjects	124.11	125			
Groups	1.10	2	0.55	0.55	ns
Subjects within groups	123.01	123	1.00		
Within subjects	75.33	252			
Session	0.64	2	0.32	1.07	ns
Groups by session	1.25	4	0.31	1.03	ns
Residual	73.44	246	0.30		

$F_{2,123}(.05) = 3.09$
$F_{2,246}(.05) = 3.04$
$F_{4,246}(.05) = 2.41$

The analysis of between-session comparisons of means for the guessing items in Test (A) was not made, since session differences were not statistically significant. Thus, the null hypothesis (H4) was upheld, so that H5-H7 did not require testing.

3.2. Test (B)

3.2.1. Feedback Items—Test (B)

The full mark of the feedback items in Test (B) was 5 points. Table 8 shows the means and standard deviations by group and session for the feedback items in Test (B). The means of correct responses for the feedback items in Test (B) are illustrated in Figure 3.

Table 9 demonstrates the results of two-way repeated-measures ANOVA for the feedback items in Test (B). The results illustrate that group differences and session differences were statistically significant ($p<.05$). Furthermore, the group by session interaction was statistically

Table 8 : Means and standard deviations by group and session for the feedback items in Test (B)

Group	n	Mean	Standard Deviation
[Pre-test]			
A	42	1.71	1.05
B	42	2.00	0.90
Z	42	1.69	0.77
[Post-test 1]			
A	42	4.55	0.85
B	42	4.71	0.50
Z	42	2.02	0.86
[Post-test 2]			
A	42	3.76	1.11
B	42	3.88	1.18
Z	42	2.00	1.07

Figure 3 : Means of correct responses for the feedback items in Test (B)

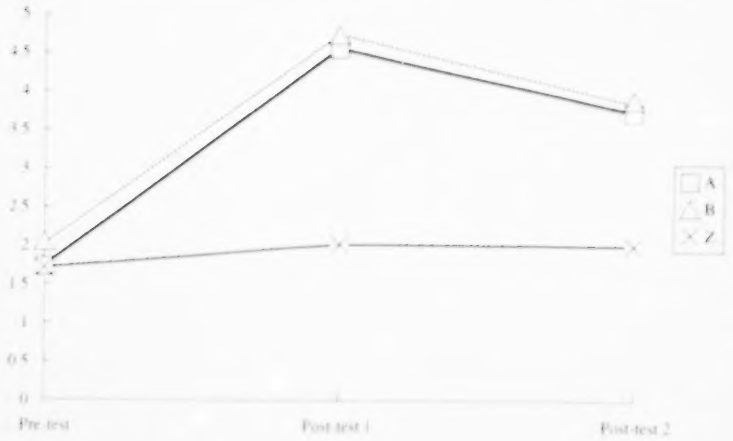

Table 9 : Results of two-way repeated-measures ANOVA for the feedback items in Test (B)

Source	SS	df	MS	F	p
Between subjects	293.81	125			
Groups	140.58	2	70.29	56.23	< .05
Subjects within groups	153.23	123	1.25		
Within subjects	527.33	252			
Session	199.03	2	99.52	134.49	< .05
Groups by session	146.01	4	36.50	49.32	< .05
Residual	182.29	246	0.74		

$F_{2,123}(.05) = 3.09$
$F_{2,246}(.05) = 3.04$
$F_{4,246}(.05) = 2.41$

Table 10 : Results of two-way repeated-measures ANOVA of the interaction effects for the feedback items in Test (B)

Source	SS	df	MS	F	p
Between subjects					
Groups in Pre-test	2.49	2	1.25	0.13	ns
Groups in Post-test 1	190.91	2	95.46	9.70	$<.05$
Groups in Post-test 2	93.19	2	46.60	4.74	$<.05$
Within cell	3630.38	369	9.84		
Within subjects					
Tests in Group A	179.73	2	89.87	121.45	$<.05$
Tests in Group B	162.39	2	81.20	109.73	$<.05$
Tests in Group Z	2.91	2	1.46	1.97	ns
Tests $\times$ $\times$	182.29	246	0.74		

$F_{2,369}(.05) = 3.04$
$F_{2,246}(.05) = 3.04$

significant. Therefore, the main effect for the group, the main effect for the session, and the interaction effect for the group by session were observed, so that groups and sessions influenced the results dependently.

The interaction effects were tested in order to determine which factors were statistically different from each other. Table 10 displays the results of the two-way repeated-measures ANOVA of the interaction effects for the feedback items in Test (B). It was confirmed that there was no statistically significant difference among groups in the Pre-test; otherwise further analyses would not have been made. Therefore, any comparative effects due to treatment were not related to prior knowledge or language ability of any one group. As shown in Table 10, groups in Post-test 1, groups in Post-test 2, tests in Group A, and tests in Group B were statistically significant ($p<.05$). Furthermore, multiple comparisons were made in order to decide which factors were different from each other, using Tukey's method. The results of between-group comparisons are shown in Tables 11 and 12, while Table 13 displays the results of

Table 11 : Between-group comparisons of means for the feedback items of Post-test 1 in Test (B)

Mean	Group	Z	A
2.02	Z		
4.55	A	*	
4.71	B	*	

*p<.05

Table 12 : Between-group comparisons of means for the feedback items of Post-test 2 in Test (B)

Mean	Group	Z	A
2.00	Z		
3.76	A	*	
3.88	B	*	

*p<.05

Table 13 : Between-session comparisons of means for the feedback items in Test (B)

Mean	Session	Pre-test	Post-test 1
[Group A]			
1.71	Pre-test		
4.55	Post-test 1	*	
3.76	Post-test 2		*
[Group B]			
2.00	Pre-test		
4.71	Post-test 1	*	
3.88	Post-test 2		*

*p<.05

between-session comparisons. As displayed in Table 11, Groups A and B were statistically different from Group Z, respectively, in Post-test 1. As shown in Table 12, Groups A and B significantly differed from Group Z, respectively, in Post-test 2. Consequently, it is concluded that treatments given to Groups A and B were more effective than no treatment (Group Z), for a long-term (one month). Therefore, the null hypothesis (H1) was rejected, thereby H2 was supported.

There was no statistically significant difference found between Groups A and B, in Post-tests 1 and 2, so that the data supported the null hypothesis (H3). As illustrated in Table 13, Post-test 1 significantly differed from the Pre-test and Post-test 2, both in Groups A and B. Accordingly, both in Tests (A) and (B), H4 was rejected, so that H5 and H7 were supported.

Table 14 : Means and standard deviations by group and session for the guessing items in Test (B)

Group	n	Mean	Standard Deviation
[Pre-test]			
A	42	1.59	1.24
B	42	1.66	1.17
Z	42	1.31	1.01
[Post-test 1]			
A	42	1.93	0.91
B	42	2.14	0.91
Z	42	1.62	1.05
[Post-test 2]			
A	42	1.74	1.24
B	42	2.12	0.89
Z	42	1.60	1.11

3.2.2. Guessing Items—Test (B)

The full mark of the guessing items in Test (B) was 5 points. Table 14 shows the means and standard deviations by each group and session for the guessing items in Test (B). The means of correct responses for the guessing items in Test (B) are illustrated in Figure 4.

Table 15 indicates the results of the two-way repeated-measures ANOVA of the guessing items in Test (B). It demonstrates that session differences were statistically significant, thereby the main effect for the session was obtained. This implies that only the sessions influenced the results independently.

Multiple comparisons (Tukey's method) were not made to determine which groups significantly differed from each other, since group differences were not of consequence. Therefore, the null hypotheses (H1 and H3) were supported.

The analysis of between-session comparisons of means for the guessing

Figure 4 : Means of correct responses for the guessing items in Test (B)

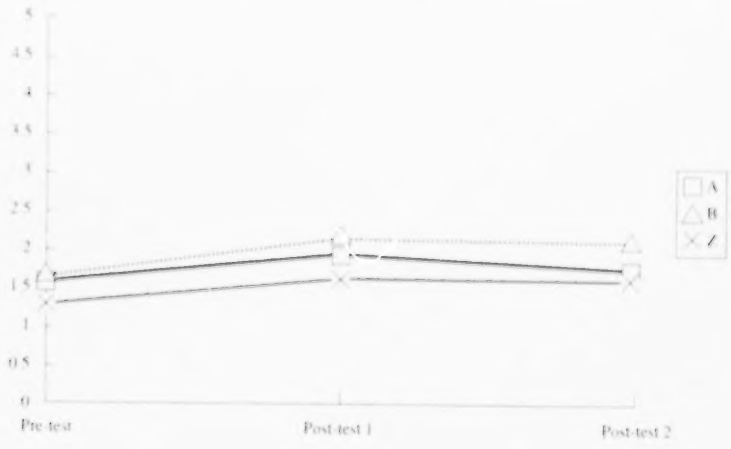

Table 15 : Results of two-way repeated-measures ANOVA for the guessing items in Test (B)

Source	SS	df	MS	F	p
Between subjects	320.29	125			
Groups	13.83	2	6.92	2.78	ns
Subjects within groups	306.46	123	2.49		
Within subjects	133.33	252			
Session	9.73	2	4.87	9.74	< .05
Groups by session	1.16	4	0.29	0.58	ns
Residual	122.44	246	0.50		

$F_{2,123}(.05) = 3.09$
$F_{2,246}(.05) = 3.04$
$F_{4,246}(.05) = 2.41$

Table 16 : Between-session comparisons of means for the guessing items in Test (B)

Mean	Session	Pre-test	Post-test 1
[Group A]			
1.59	Pre-test		
1.93	Post-test 1		
1.74	Post-test 2		
[Group B]			
1.66	Pre-test		
2.14	Post-test 1	*	
2.12	Post-test 2		•
[Group Z]			
1.31	Pre-test		
1.62	Post-test 1		
1.60	Post-test 2		

* $p < .05$

items in Test (B) was made in order to determine which sessions were significantly different from each other. Table 16 shows the between-session comparisons of means for the guessing items in Test (B). As illustrated in Table 16, Post-test 1 significantly differed from the Pre-test in Group B. Hence, only H5 was upheld out of three hypotheses (H5–H7).

4. DISCUSSION

The following hypotheses were tested in this study:

H1: There would be no statistically significant difference in the accuracy of responses between the experimental groups and the control group.

H2: The experimental groups would outperform the control group in Post-test 1 and Post-test 2.

H3: There would be no statistically significant difference in the accuracy of responses between the experimental groups.

H4: There would be no statistically significant difference in the accuracy of responses among the Pre-test, Post-test 1, and Post-test 2, in the experimental groups.

H5: The experimental groups would respond significantly better in Post-test 1 than in the Pre-test.

H6: The experimental groups would respond significantly better in Post-test 2 than in the Pre-test.

H7: The experimental groups would respond significantly better in Post-test 1 than in Post-test 2.

4.1. Feedback Items

H1–H3: In Test (A), which required the subjects to choose the appropriate response in each situation, the experimental groups (Groups A and B) outperformed the control group only in Post-test 1. This demonstrates that the subjects in Groups A and B got statistically higher scores in Post-test 1 than in the Pre-test, so that the treatments in Groups A and B had an immediate impact on learners' pragmatic knowledge. That is,

the learning effect continued in the short term.

In Test (B), which asked the subjects to write conversational implicatures, H2 was upheld in that Groups A and B did better than Group Z in both Post-tests 1 and 2. This result shows that the learning effect was maintained by the provision of explanations (Group A) and by consciousness-raising tasks (Group B). Therefore, in light of the results of Tests (A) and (B), it is concluded that explanations of rules concerning conversational implicatures and consciousness-raising tasks were both beneficial, at least in the short term; however, it was found in both Tests (A) and (B) that there was no statistically significant difference in the accuracy of responses between Groups A and B, in Post-tests 1 and 2.

These interesting results demonstrate that all the experimental groups generated significantly better responses, at least temporarily. However, it is necessary in the scope of this research to examine what kind of instruction may be necessary to hold long-term learning durability. The evidence suggests that teachers should be aware that EFL instruction may include explicit explanations of rules and consciousness-raising tasks.

H4-H7: In Test (A), the subjects of Group B performed better in Post-test 1 than in the Pre-test. On the contrary, it was found that no tests significantly differed from each other in Group A, even though ANOVA revealed a statistically significant difference among the tests in Group A.

In Test (B), scores in Post-test 1 were found statistically higher than those in the Pre-test and Post-test 2, in both Groups A and B.

In summary, it was found that the subjects performing consciousness-raising tasks did significantly better in Post-test 1 than in the Pre-test, regardless of tests. This phenomenon may imply that it would be advantageous for learners to process language on their own through consciousness-raising tasks.

BEST COPY AVAILABLE

4.2. Guessing Items

H1-H3: In Tests (A) and (B), the null hypothesis (H1) was upheld, so that there was no statistically significant difference in the accuracy of responses between the experimental groups (A and B) and the control group (Z). This result clearly indicates that the guessing items failed to obtain significant results. The data prove that the subjects did not extract the expected pragmatic generalizations from the treatment that they were applying to the new items. It may be possible that "the limited amount of time and exposure to the linguistic system" (Nayak et al. 1990) influenced the results.

Furthermore, no statistically significant difference was found between Groups A and B, in Tests (A) and (B).

H4-H7: In Test (A), there was no statistically significant difference in accuracy among the responses in the Pre-test, Post-test 1, and Post-test 2, in the experimental groups (A and B). In Test (B), however, the null hypothesis (H4) was not supported, thereby H5 was upheld in that Group B got significantly higher scores in Post-test 1 than in the Pre-test. The reason for this phenomenon may be that there existed the effect of Pre-test experiences, that is, Post-test 1 was the second time in which the same test was given to the subjects.

5. CONCLUSION

The following major findings and pedagogical implications emerge from this classroom research:

(1) All the experimental groups (A and B) generated significantly better responses in the feedback items than the control group (Z), at least temporarily. This evidence suggests that teachers keep in mind that EFL instruction in teaching pragmatic knowledge may include explicit explanations of rules and consciousness-raising tasks.

(2) The subjects performing consciousness-raising tasks (Group B) did significantly improve in Post-test 1 over the results found in the Pre-test,

in the feedback items. It seems that it was advantageous for the learners to process language on their own in consciousness-raising tasks.

(3) No subjects extracted the expected pragmatic generalizations from the treatment that they were applying to the new items. Teachers should notice that the amount of time and exposure to the pragmatic system may be a crucial factor to induction.

(4) The subjects in Group B got significantly higher scores in the guessing items of Post-test 1 than in the Pre-test. There may exist the effect of Pre-test experiences, that is, Post-test 1 was the second time in which the same test was given to the subjects.

In summary, the results of this research confirmed the assumption that it is effective to teach conversational implicature to EFL learners in Japan.

The limitations of this research are that (1) the number of tests was only 2 all written tests, and (2) the number of subjects in each group was only 42. Thus, this research should be viewed as a pilot study that was first conducted in Japan.

It is interesting to note that No.6 in Test (B) proved the most difficult item of all; only 7 (2.8%) of 246 subjects got the correct response. This test item includes the conversational implicature breaking the maxim of manner. The difficulty of this item may be due to the idea that the subjects found it difficult to understand the relatively long sentence even at a literal level. At any rate, the types of conversational implicatures which may be difficult were not investigated, since each type contained only one or two examples in each test. This topic needs further investigation.

Further research may include the following experimental groups: (a) groups without Pre-tests should be added to examine the effect of Pre-test experiences, and (b) groups that were given both explanations of rules and opportunities of doing consciousness-raising tasks should be included. The research questions, such as (1) what type of instruction should

be provided for learners to keep long-term retention of pragmatic knowledge, and (2) in what types of linguistic items learners may induce generalizations from the treatment, may be investigated.

NOTES

* I would like to thank Sara M. Luna (Lakeland College, WI, USA) and *The IRLT Bulletin* reviewers for their valuable comments that contributed to the improvements of this research.

1. Holmes (1995) divided pragmatic competence into 'pragmalinguistic competence' and 'sociopragmatic competence,' based on Thomas (1983). The former refers to the knowledge underlying a person's ability to correctly interpret the intended illocutionary force of an utterance, or to encode illocutionary force appropriately. The latter means the knowledge underlying a person's ability to express and interpret meaning appropriately in a range of different social situations. This research tries to investigate the teachability of pragmalinguistic competence.
2. In Kubota's (1994) study, only one group given modeling and implicit feedback out of four experimental groups performed better than the control group.
3. The results of Carroll et al. (1992) showed that the advanced groups in the experimental groups gained learning between the feedback session and the second recall session (one week later). Kubota (1994) found that all the four experimental groups did better in the Post-test 1 (immediate post-test) than in the Pre-test. Furthermore, Kubota (1995) reported that the group receiving instruction (the Garden Path technique) performed better in Post-test 1 (immediate post-test) than in the Pre-test and Post-test 2 (one week after treatment), on two tests given. However, these three studies do not guarantee an advantage of instruction over a long period of time.
4. "Is the Pope Catholic?" is the sentence from which this implicature gets its name. A question to which the answer is *Yes* may be the basis of a POPE Q implicature, as in No.2 of Test (A). However, the answer *No* can be a response less frequently, as in No.6 of Test (A).
5. One discussion group includes 5 subjects, in each one of Groups A and Groups B.
6. The present researcher's observation by circulating in class revealed that approximately three-fourths of group discussions were conducted in the target language.

REFERENCES

Billmyer, K. "'I really like your life style': ESL learners learning how to compli-

ment." *Penn Working Papers in Educational Linguistics*, 6, 2 (1990), 31–48.

Bouton, L.F. "A cross-cultural study of ability to interpret implicatures in English." *World Englishes*, 7 (1988), 183–196.

———. "The effective use of implicature in English: why and how it should be taught in the ESL classroom." In Bouton, L.F. and Y. Kachru, eds., (1990), 43–52.

———. "Conversational implicature in a second language: learned slowly when not deliberately taught." *Journal of Pragmatics*, 22, 2 (1994), 157–167.

Bouton, L.F. and Y. Kachru, eds. *Pragmatics and Language Learning*. (Monograph series) Urbana, IL: University of Illinois, 1990.

Carroll, S., Y. Roberge, and M. Swain. "The role of feedback in adult second language acquisition: error correction and morphological generalizations." *Applied Linguistics*, 13, 3 (1992), 173–198.

Carroll, S. and M. Swain. "Explicit and implicit negative feedback: an empirical study of the learning of linguistic generalizations." *Studies in Second Language Acquisition*, 15, 3 (1993), 357–386.

Cohen, A.D. and E. Olshtain. "The learning of complex speech act behavior." Paper presented at 22nd annual TESOL convention, Chicago, 1988.

Cole, P. and J. Morgan, eds. *Syntax and Semantics*, *Vol.3*. New York: Academic Press, 1975.

Ellis, R. *The Study of Second Language Acquisition*. Oxford: Oxford University Press, 1994.

Fujii, T., ed. *Eigo no Sekai* (The World of English). Kyoto: Shohwadou, 1992.

Fujimura, T., Y. Kato, M. Leong, and R. Uehara, eds. *Proceedings of the 5th Conference on Second Language Research in Japan*. Niigata: International University of Japan, 1994.

Grice, H.P. "Logic and conversation." In Cole, P. and J. Morgan, eds., (1975), 41–58.

Holmes, J. "Sociolinguistic competence, gender and second language learning." Temple University Japan Seminar. Tokyo: Temple University Japan, Feb. 11–12, 1995.

Kasper, G. "Interlanguage pragmatics and second language acquisition." In Fujimura, T. et al., eds., (1994), 25–39.

Kinoshita, H. *Imiron to Goyouron* (Semantics and Pragmatics). In Fujii, T. (ed.), (1992), 127–153.

Kubota, M. "The role of negative feedback on the acquisition of the English dative alternation by Japanese college students of EFL." *The IRLT Bulletin*, 8 (1994), 1–36. Tokyo, Japan: Institute for Research in Language Teaching.

———. "The Garden Path technique: is it really effective?" *Working Papers of Chofu*

BEST COPY AVAILABLE

Gakuen Women's Junior College, 27 (1995), 21-48. Kawasaki, Japan: Chofu Gakuen Women's Junior College.

Nayak, N., N. Hansen, N. Krueger, and B. McLaughlin. "Language-learning strategies in monolingual and multilingual subjects." *Language Learning*, 40, 2 (1990), 221-244.

Thomas, J. "Cross-cultural pragmatic failure." *Applied Linguistics*, 4, 2 (1983), 91-112.

Wilder-Bassett, M. *Improving Pragmatic Aspects of Learners' Interlanguage*. Tuebingen: Narr, 1984.

APPENDIX: TEST (cf. Grice 1975, Bouton 1994)

Test (A): Choose the appropriate response in each situation.

1. Two friends, Maria and Tony, are talking about what happened the night before. They had dinner with Sandy, a friend of theirs, in a little town just outside Philadelphia. Then, after dinner, Sandy left. Now this morning, Maria and Tony are trying to figure out what Sandy did after he left them.

   Maria: Hey, I hear Sandy went to Philadelphia and stole a car after he left us last night.

   Tony: Not exactly. He stole a car and went to Philadelphia.

   Maria: Are you sure? That's not the way I heard it.

   What actually happened is that Sandy stole the car in Philadelphia itself. In that case, which of the two friends has the right story—Maria or Tony?

   (a) Maria.

   (b) Tony.

   (c) Both are right. Since they are both saying essentially the same thing, they really have nothing to argue about.

   (d) Neither of them has the story right.

2. Two roommates are talking about their plans for the summer.

   Fran: My mother wants me to stay home for a while, so I can be there when our relatives come to visit us at the beach.

   Joan: Do you have a lot of relatives?

   Fran: Is the sky blue?

   How can we best interpret Fran's question?

   (a) Fran thinks her relatives are all blue.

   (b) Fran is new to the area and is trying to find out what the summers are like.

   (c) Fran has a lot of relatives.

   (d) Fran is trying to change the subject; she doesn't want to talk about her relatives.

3. Frank wanted to know what time it was, but he didn't have a watch.

Frank: What time is it, Helen?

Helen: The postman has been here.

Frank: Okay. Thanks.

(a) She is telling him approximately what time it is by telling him that the postman has already been here.

(b) By changing the subject, Helen is telling Frank that she doesn't know what time it is.

(c) She thinks that Frank should stop what he is doing and read his mail.

(d) Frank will not be able to interpret any message from what Helen says, since she did not answer the question.

4. Mr. Rose was murdered at his house. A police officer conducted the following interview:

Police officer: Mrs. Rose, what did you see after you heard the shot?

Mrs. Rose : I saw a man running out of the house and then changing his clothes once outside.

Rose's son : I thought the man first changed his clothes and ran out of the house.

Mrs. Rose : I don't remember it that way.

The true information was that a murderer changed his clothes inside the house. Who said the correct story?

(a) Mrs. Rose.

(b) Mrs. Rose's son.

(c) Both are right.

(d) Neither of them is right.

5. Ken bought a new car and his friend, Charles, came to see it. Charles drove it around for an hour near Ken's house.

Ken: What do you think of this new car?

Charles: Well, the color's fine, but...

Ken: Thanks.

How can we interpret Ken's response?

(a) Ken appreciates Charles for driving Ken's car.

(b) Ken is happy to have this fantastic car.

(c) Ken feels happy, because Charles praises the color of the car.

(d) Ken does not want to hear any other criticism from Charles.

6. Two friends are working together in a construction site. After Bill finished his job, he asked John, who is totally bald.

Bill: Can I give you a hand?

John: Do I need a haircut?

How can we best interpret John?

(a) John interprets that Bill is a good hairdresser.

(b) John thinks that he needs to have his hair cut.

(c) John does not need any help.

(d) John is trying to change the subject; he does not talk with Bill.

Test (B): Write what the speaker in the boldface type is trying to say.

1. John is driving his car with Mike.

John: I am out of gas.

Mike: **There is a gas station round the corner.**

•Mike: _____

2. A professor makes the following recommendation about Ken who wants to study English in America next year:

Ken's command of Japanese is excellent, and his class attendance has been regular.

•professor: _____

3. Mary told a lie to Tim.

Tim: **Mary is a nice friend.**

•Tim: _____

4. At a tea ceremony:

A: Mr. Johnson shouldn't have come to the party today. Who invited him?

B: **The weather has been wonderful today, hasn't it?**

•B: _____

5. A: Smith doesn't seem to have a girlfriend these days.

B: **He has been visiting New York lately.**

•B: _____

6. A: Kate sang 'Endless Love.'

B: Yeah, **she produced a series of sounds that corresponded closely with the score of 'Endless Love.'**

•B: _____

7. A: Where does Max live?

B: **Somewhere in the south of France.**

•B: _____

8. Customer: **When is the Oxford train?**
Official: The train to Oxford?
Customer: Yes.
Official: 3:15.
→Customer: _____

9. Mike destroyed the furniture completely when he came home from a party. Mike's wife said the next morning:
Mike's wife: **You were a little drunk last night.**
→Mike's wife: _____

10. Kathy: Can you answer the phone?
Bob: **I'm in the bath.**
→Bob: _____

(調布学園女子短期大学)